American Public University System
American Military University | American Public University

DATS311

Course Summary

Course : DATS311 **Title :** Intermediate Data Science
Length of Course : 8
Prerequisites : DATS301 Analytical Methods II **Credit Hours : **3

Description

Course Description:

This course continues to expand the knowledge, skills and abilities of students by two paths; first through the design of experiments required to acquire specified data, and second through the use of carefully designed experiments to establish causal effects. Students will take a deep dive into the differences between correlation and causality. They will learn the critical thinking skills required to assert the reliability of data acquired.

Course Scope:

This course is intended for students seriously pursuing an undergraduate degree in data science. This course assumes a fairly high level of mathematics knowledge in calculus as well as probability and statistics. Causality is one of the main topics in this course. Students will be introduced to the concept of the design of experiments. In addition, this course provides students with knowledge about how to handle data acquired in a variety of experimental situations, e.g. randomization, blocking or with a variety of contrasts or treatments.

Objectives

CO1:   Describe the differences between correlation and causation.
CO2:   Build basic, structural causal models.
CO3:   Implement causality in statistical analyses.
CO4:   Explain how experimentation affects the basic causal inference problem.
CO5:   Describe the differences between laboratory and field experiments.
CO6:   Demonstrate why control groups and placebos are both important.
CO7:   Apply the power of randomization and "as if" random research designs in dealing with confounding factors.
CO8:   Design, implement and analyze a field experiment to evaluate causation.

Outline

Week 1: Correlation versus Causation

Learning Objectives:

CO1: Describe the differences between correlation and causation.

Reading(s)

1. Online readings as laid out in Week 1 (URLs/links provided)
2. Pearl, **Causal Inference in Statistics: A Primer** read Section 1.1 and review the material in Section 1.3

Assignment(s)

Week 1 Welcome Discussion
Week 1 Discussion: Why isn't this causation?
Assignment #1: Correlation versus Causation

Week 2: Simpson's Paradox

Learning Objectives:

CO2: Explain how experimentation affects the basic causal inference problem.
CO5: Apply the power of randomization and "as if" random research designs in dealing with confounding factors.

Reading(s)

1. Watch the online video listed in the Learning Material block (URL/link provided)
2. Read "Race, COVID Mortality, and Simpson's Paradox" by Dana Mackenzie (URL/link provided in the Learning Material block)
3. Pearl, **Causal Inference in Statistics: A Primer** read Section 1.2 on Simpson's Paradox

Assignment(s)

Week 2 Discussion: Is this Causation?
Assignment #2: Simpson's Paradox

Week 3: Challenges, Variables that Affect Responses

Learning Objectives:

CO4: Demonstrate why control groups and placebos are both important.
CO5: Apply the power of randomization and "as if" random research designs in dealing with confounding factors.

Reading(s)

1. Read "Mediation" by David Kenny (URL/link provided)
2. Read "Moderation" by David Kenny (URL/link provided)
3. Read "Assessing Bias: the importance of considering confounding" by Skelly, Dettori and Brodt (provided in the course files)

Assignment(s)

Week 3 Discussion: When "unknown" variables are present?
Assignment #3: Correlation versus Causation (Long assignment in 3 part over multiple weeks)
Assignment #4: Mediating, Moderating and Confounding Variables

Week 4: Causal Models

Learning Objectives:

CO1: Describe the differences between correlation and causation.

CO4: Explain how experimentation affects the basic causal inference problem.

CO7: Apply the power of randomization and "as if" random research designs in dealing with confounding factors.

CO8: Design, implement and analyze a field experiment.

Reading(s)

Pearl, **Causality** Chapter 1, Section 1.2

Pearl, **Causal Inference in Statistics: A Primer**, Chapter 1, Sections 1.4 and 1.5

Assignment(s)

Week 4 Discussion: When to moderate or when to mediate?
Assignment #5: Causal Models
Quiz #1

Week 5: Design of Experiments

Learning Objectives:

CO5: Describe the differences between laboratory and field experiments.

CO6: Demonstrate why control groups and placebos are both important.

CO7: Apply the power of randomization and "as if" random research designs in dealing with confounding factors.

CO8: Design, implement and analyze a field experiment.

Reading(s)

Antony, **Design of Experiments for Engineers and Scientists**, Chapters 2, 3, and 4
The National Institute of Standards (NIST) **Engineering Statistics Handbook**, Section 5.1 (online, URL/link provided)

Assignment(s)

Week 5 Discussion: What is the "curse of dimensionality"?
Assignment #6: Correlation versus Causation (Long assignment over multiple weeks continued)
Course Project Proposal

Week 6: Acquiring and Analyzing Experimental Data

Learning Objectives:

CO4: Explain how experimentation affects the basic causal inference problem.

CO6: Demonstrate why control groups and placebos are both important.

CO7: Apply the power of randomization and "as if" random research designs in dealing with confounding factors.

CO8: Design, implement and analyze a field experiment.

Reading(s)

Antony, **Design of Experiments for Engineers and Scientists**, Chapters 6 and 8

The National Institute of Standards (NIST) **Engineering Statistics Handbook**, remaining sections in Chapter 5 (online, URL/link provided)

Assignment(s)

Week 6 Discussion: How much data should you acquire?
Assignment #7: Regression
Quiz #2

Week 7: Causality, Linear Regression and Contingency Tables

Learning Objectives:

CO2: Build basic, structural causal models.

CO3: Implement causality in statistical analyses.

Reading(s)

Agresti, **Introduction to Categorical Data Analysis**, (textbook from DATS301) Review Chapters 1, 2, and 3

Best and Wolf, **The SAGE Handbook of Regression Analysis and Causal Inference**, Chapters 2, 4, and 5 (these chapters review linear regression)

Assignment(s)

Week 7 Discussion: What are the steps in a good research project?
Assignment #8: Correlation versus Causation (conclusion of long assignment over multiple weeks)

Week 8: Causality and Logistic Regression

Learning Objectives:

CO2: Build basic, structural causal models.

CO3: Implement causality in statistical analyses.

Reading(s)

Agresti, **Introduction to Categorical Data Analysis**, (textbook from DATS301) Review Chapters 4 and 5

Best and Wolf, **The SAGE Handbook of Regression Analysis and Causal Inference**, Chapters 8 and 9 (these chapters review logistic regression)

Assignment(s)

Week 8 discussion: How do you decide what "type" of data to acquire?
Course Project Final Written Report and Presentation
Quiz #3

Evaluation

Grading will be based on:

- Eight Discussions (20%)

- Eight Assignments (30%) (NOTE: some assignments are longer than others. To compensate you will have multiple weeks to complete the longer assignments.)

- Three Quizzes (20%)

- And, a Course Project (30%)

[Detailed directions on each of these assessments are provided in MyClassroom.]

Materials

All required and optional course materials are provided in the course (as outlined in the weekly lessons above), in course files, through the Trefry Library or online via URLs/links.

Course Guidelines

Citation and Reference Style
Attention Please: Students will follow the APA Format as the sole citation and reference style used in written work submitted as part of coursework to the University. Assignments completed in a narrative essay or composition format must follow the citation style cited in the APA Format.

Tutoring
Tutor.com offers online homework help and learning resources by connecting students to certified tutors for one-on-one help. AMU and APU students are eligible for 10 free hours* of tutoring provided by APUS. Tutors are available 24/7 unless otherwise noted. Tutor.com also has a SkillCenter Resource Library offering educational resources, worksheets, videos, websites and career help. Accessing these resources does not count against tutoring hours and is also available 24/7. Please visit the APUS Library and search for 'Tutor' to create an account.

Late Assignments
Students are expected to submit classroom assignments by the posted due date and to complete the course according to the published class schedule. The due date for each assignment is listed under each Assignment. Generally speaking, late work may result in a deduction up to 15% of the grade for each day late, not to exceed 5 days. As a working adult I know your time is limited and often out of your control. Faculty may be more flexible if they know ahead of time of any potential late assignments.

Turn It In
Faculty may require assignments be submitted to Turnitin.com. Turnitin.com will analyze a paper and report instances of potential plagiarism for the student to edit before submitting it for a grade. In some cases professors may require students to use Turnitin.com. This is automatically processed through the Assignments area of the course.

Academic Dishonesty
Academic Dishonesty incorporates more than plagiarism, which is using the work of others without citation. Academic dishonesty includes any use of content purchased or retrieved from web services such as CourseHero.com. Additionally, allowing your work to be placed on such web services is academic dishonesty, as it

is enabling the dishonesty of others. The copy and pasting of content from any web page, without citation as a direct quote, is academic dishonesty. When in doubt, do not copy/paste, and always cite.

Submission Guidelines
Some assignments may have very specific requirements for formatting (such as font, margins, etc) and submission file type (such as .docx, .pdf, etc) See the assignment instructions for details. In general, standard file types such as those associated with Microsoft Office are preferred, unless otherwise specified.

Disclaimer Statement
Course content may vary from the outline to meet the needs of this particular group.

Communicating on the Discussion
Discussions are the heart of the interaction in this course. The more engaged and lively the exchanges, the more interesting and fun the course will be. Only substantive comments will receive credit. Although there is a final posting time after which the instructor will grade comments, it is not sufficient to wait until the last day to contribute your comments/questions on the discussion. The purpose of the discussions is to actively participate in an on-going discussion about the assigned content. "Substantive" means comments that contribute something new and hopefully important to the discussion. Thus a message that simply says "I agree" is not substantive. A substantive comment contributes a new idea or perspective, a good follow-up question to a point made, offers a response to a question, provides an example or illustration of a key point, points out an inconsistency in an argument, etc. As a class, if we run into conflicting viewpoints, we must respect each individual's own opinion. Hateful and hurtful comments towards other individuals, students, groups, peoples, and/or societies will not be tolerated.

University Policies

Student Handbook
Drop/Withdrawal policy
Extension Requests
Academic Probation
Appeals
Disability Accommodations

The mission of American Public University System is to provide high quality higher education with emphasis on educating the nation's military and public service communities by offering respected, relevant, accessible, affordable, and student-focused online programs that prepare students for service and leadership in a diverse, global society.